GEOMETRIC STRUCTURE-PRESERVING OPTIMAL CONTROL OF THE RIGID BODY

ANTHONY M. BLOCH, ISLAM I. HUSSEIN, MELVIN LEOK, AND AMIT K. SANYAL

ABSTRACT. In this paper we study a discrete variational optimal control problem for the rigid body. The cost to be minimized is the external torque applied to move the rigid body from an initial condition to a pre-specified terminal condition. Instead of discretizing the equations of motion, we use the discrete equations obtained from the discrete Lagrange–d'Alembert principle, a process that better approximates the equations of motion. Within the discrete-time setting, these two approaches are not equivalent in general. The kinematics are discretized using a natural Lie-algebraic formulation that guarantees that the flow remains on the Lie group $\mathsf{SO}(3)$ and its algebra $\mathfrak{so}(3)$. We use Lagrange's method for constrained problems in the calculus of variations to derive the discrete-time necessary conditions. We give a numerical example for a three-dimensional rigid body maneuver.

1. INTRODUCTION

This paper deals with a structure-preserving computational approach to the optimal control problem of minimizing the control effort necessary to perform an attitude transfer from an initial state to a prescribed final state, in the absence of a potential field. The configuration of the rigid body is given by the rotation matrix from the body frame to the spatial frame, which is an element of the group of orientation-preserving isometries in $\mathbb{R}^3$. The state of the rigid body is described by the rotation matrix and its angular velocity.

To motivate the computational approach we adopt in the discrete-time case, we first revisit the variational continuous-time optimal control problem. The continuous-time extremal solutions to this optimal control problem have certain special features, since they arise from variational principles. General numerical integration methods, including the popular Runge-Kutta schemes, typically preserve neither first integrals nor the characteristics of the configuration space. Geometric integrators are the class of numerical integration schemes that preserve such properties, and a good survey can be found in [5]. Techniques particular to Hamiltonian systems are also discussed in [16] and [24].

Our approach to discretizing the optimal control problem is in contrast to traditional techniques such as collocation, wherein the continuous equations of motion are imposed as constraints at a set of collocation points. In our approach, modeled after [11], the discrete equations of motion are derived from a discrete variational principle, and this induces constraints on the configuration at each discrete time step.

This approach yields discrete dynamics that are more faithful to the continuous equations of motion, and consequently yields more accurate solutions to the optimal

control problem that is being approximated. This feature is extremely important in computing accurate (sub)optimal trajectories for long-term spacecraft attitude maneuvers. For example, in [9], the authors propose an imaging spacecraft formation design that requires a continuous attitude maneuver over a period of 77 days in a low Earth orbit. Hence, attitude maneuver has to be very accurate to meet tight imaging constraints over long time ranges.

While the discrete optimal control method presented here is illustrated using the Lie group $\mathsf{SO}(3)$ of rotation matrices, and its corresponding Lie algebra $\mathfrak{so}(3)$ of skew-symmetric matrices, we have derived the method with sufficient generality to address the problem of optimal control on arbitrary Lie groups with the drift vector field given by geodesic flow on the group, and it therefore widely applicable. For example, in inter-planetary orbit transfers (see, for example, [1]), one is interested in computing optimal or suboptimal trajectories on the group of rigid body motions $\mathsf{SE}(3)$ with a high degree of accuracy. Similar requirements also apply to the control of quantum systems. For example, efficient construction of quantum gates is a problem on the unitary Lie group $\mathsf{SU}(N)$. This is an optimal control problem, where one wishes to steer the identity operator to the desired unitary operator (see, for example, [13] and [23]).

Moreover, an important feature of the way we discretize the optimal control problem is that it is $\mathsf{SO}(3)$-equivariant. The $\mathsf{SO}(3)$-equivariance of our numerical method is desirable, since it ensures that our results do not depend on the choice of coordinates and coordinate frames. This is in contrast to methods based on coordinatizing the rotation group using quaternions, (modified) Rodrigues parameters, and Euler angles, as given in the survey [26]. Even if the optimal cost function is $\mathsf{SO}(3)$-invariant, as in [25], the use of generalized coordinates imposes constraints on the attitude kinematics.

For the purpose of numerical simulation, the corresponding discrete optimal control problem is posed on the discrete state space as a two stage discrete variational problem. In the first step, we derive the discrete dynamics for the rigid body in the context of discrete variational mechanics [20]. This is achieved by considering the discrete Lagrange–d'Alembert variational principle [12] in combination with essential ideas from Lie group methods [10], which yields a Lie group variational integrator [17]. This integrator explicitly preserves the Lie group structure of the configuration space, and is similar to the integrators introduced in [14] for a rigid body in an external field, and in [15] for full body dynamics. These discrete equations are then imposed as constraints to be satisfied by the extremal solutions to the discrete optimal control problem, and we obtain the discrete extremal solutions in terms of the given terminal states.

The paper is organized as follows. As motivation, in Section 2, we study the minimum control effort optimal control problem in continuous-time. In Section 3, we study the corresponding discrete-time optimal control problem. In Section 3.1 we state the optimal control problem and describe our approach. In Section 3.2, we derive the discrete-time equations of motion for the rigid body starting with the discrete Lagrange–d'Alembert principle. These equations are used in Section 3.3 to obtain the solution to the discrete optimal control problem. In Section 5, we describe an algorithm for solving the general nonlinear, implicit necessary conditions for $\mathsf{SO}(3)$ and give numerical examples for rest-to-rest and slew-up spacecraft maneuvers.

2. Continuous-Time Results

2.1. **Problem Formulation.** In this paper, the natural pairing between $\mathfrak{so}^*(3)$ and $\mathfrak{so}(3)$ is denoted by $\langle \cdot, \cdot \rangle$. Let $\ll \cdot, \cdot \gg$ and $\ll \cdot, \cdot \gg_*$ denote the standard (induced by the Killing form) inner product on $\mathfrak{so}(3)$ and $\mathfrak{so}^*(3)$, respectively. The inner product $\ll \cdot, \cdot \gg_*$ is naturally induced from the standard norm $\ll \boldsymbol{\xi}, \boldsymbol{\omega} \gg = -\frac{1}{2}\text{Tr}(\boldsymbol{\xi}^T \boldsymbol{\omega})$, for all $\boldsymbol{\xi}, \boldsymbol{\omega} \in \mathfrak{so}(3)$, through

$$\ll \boldsymbol{\eta}, \boldsymbol{\varphi} \gg_* = \langle \boldsymbol{\eta}, \boldsymbol{\varphi}^\sharp \rangle = \langle \boldsymbol{\eta}, \boldsymbol{\omega} \rangle = \left\langle \boldsymbol{\xi}^\flat, \boldsymbol{\omega} \right\rangle$$
$$(1) \qquad\qquad = \ll \boldsymbol{\xi}, \boldsymbol{\omega} \gg,$$

where $\boldsymbol{\varphi} = \boldsymbol{\omega}^\flat \in \mathfrak{so}^*(3)$ and $\boldsymbol{\eta} = \boldsymbol{\xi}^\flat \in \mathfrak{so}^*(3)$, with $\boldsymbol{\xi}, \boldsymbol{\omega} \in \mathfrak{so}(3)$ and $\flat$ and $\sharp$ are the musical isomorphisms with respect to the standard metric $\ll \cdot, \cdot \gg$. On $\mathfrak{so}(3)$, these isomorphisms correspond to the transpose operation. That is, we have $\boldsymbol{\varphi} = \boldsymbol{\omega}^\text{T}$ and $\boldsymbol{\eta} = \boldsymbol{\xi}^\text{T}$.

Let $\mathbf{J} : \mathfrak{so}(3) \to \mathfrak{so}^*(3)$ be the positive definite inertia operator. It can be shown that

$$(2) \qquad\qquad \langle \mathbf{J}(\boldsymbol{\xi}), \boldsymbol{\omega} \rangle = \langle \mathbf{J}(\boldsymbol{\omega}), \boldsymbol{\xi} \rangle .$$

On $\mathfrak{so}(3)$, $\mathbf{J}$ is given by $\mathbf{J}(\boldsymbol{\xi}) = J\boldsymbol{\xi} + \boldsymbol{\xi}J$, where J is a positive definite symmetric matrix (see, for example, [3, 8]). Moreover, we also have $\mathbf{J}(\boldsymbol{\eta}^\sharp)^\sharp = (J\boldsymbol{\eta}^\text{T} + \boldsymbol{\eta}^\text{T}J)^\text{T} = \mathbf{J}(\boldsymbol{\eta})$, which is an abuse of notation since $\boldsymbol{\eta} \in \mathfrak{so}^*(3)$. For the sake of generality and mathematical precision we will use the general definitions, though it helps to keep the above identifications for $\mathfrak{so}(3)$ in mind.

In this section we review some continuous-time optimal control results using a simple optimal control example on $\mathsf{SO}(3)$. The problem we consider is that of minimizing the norm squared of the control torque $\boldsymbol{\tau} \in \mathfrak{so}^*(3)$ applied to rotate a rigid body subject to the Lagrange–d'Alembert principle for the rigid body[1] whose configuration is given by $\mathbf{R} \in \mathsf{SO}(3)$ and body angular velocity is given by $\boldsymbol{\Omega} \in \mathfrak{so}(3)$. We require that the system evolve from an initial state $(\mathbf{R}_0, \boldsymbol{\Omega}_0)$ to a final state $(\mathbf{R}_T, \boldsymbol{\Omega}_T)$ at a fixed terminal time T.

Before proceeding with a statement of the optimal control problem, we first define variations of the rigid body configuration $\mathbf{R}$ and its velocity $\boldsymbol{\Omega}$. Let $\mathbf{W}(t) \in \mathfrak{so}(3)$ be the variation vector field associated with a curve $\mathbf{R}(t)$ on $\mathsf{SO}(3)$ [2, 28]. The vector field $\mathbf{W}(t)$ satisfies

$$\boldsymbol{\delta}\mathbf{R}(t) = \mathbf{R}\mathbf{W} \in \mathsf{T}_{\mathbf{R}(t)}\mathsf{SO}(3), \ \mathbf{W}(0) = \mathbf{0}, \ \mathbf{W}(T) = \mathbf{0},$$

where $\boldsymbol{\delta}\mathbf{R}$ is defined by $\boldsymbol{\delta}\mathbf{R}(t) = \partial \mathbf{R}_\epsilon(t)/\partial \epsilon \big|_{\epsilon=0}$, with $\mathbf{R}_\epsilon(t) := \mathbf{R}(t, \epsilon)$ is the variation of the curve $\mathbf{R}(t)$ that satisfies $\mathbf{R}(t, 0) = \mathbf{R}(t)$. The variation in the velocity vector field is denoted $\boldsymbol{\delta}\boldsymbol{\Omega}$. For a deeper understanding of variations of general vector fields, see for example the treatment in [7].

We now state the minimum control effort optimal control problem.

---

[1]This is equivalent to constraining the problem to satisfy the rigid body equations of motion given by equations (7). However, for the sake of generality that will be appreciated in the discrete-time problem, we choose to treat the Lagrange–d'Alembert principle as the constraint as opposed to the rigid body equations of motion. Both are equivalent in the continuous-time case but are generally not equivalent in the discrete-time case.

Problem 2.1. *Minimize*

$$(3) \qquad \mathcal{J} \;=\; \frac{1}{2} \int_0^T \ll \boldsymbol{\tau}, \boldsymbol{\tau} \gg_* \mathrm{d}t$$

subject to

(1) *satisfying* Lagrange–d'Alembert principle*:*

$$(4) \qquad \boldsymbol{\delta} \int_0^T \frac{1}{2} \langle \mathbf{J}\left(\boldsymbol{\Omega}\right), \boldsymbol{\Omega} \rangle \, \mathrm{d}t + \int_0^T \langle \boldsymbol{\tau}, \mathbf{W} \rangle \, \mathrm{d}t = 0,$$

for a variation vector field $\mathbf{W}(t)$, *and subject to* $\dot{\mathbf{R}} = \mathbf{R}\boldsymbol{\Omega}$,

(2) *and the* boundary conditions

$$(5) \qquad \begin{aligned} \mathbf{R}(0) &= \mathbf{R}_0, \; \boldsymbol{\Omega}(0) = \boldsymbol{\Omega}_0, \\ \mathbf{R}(T) &= \mathbf{R}_T, \; \boldsymbol{\Omega}(T) = \boldsymbol{\Omega}_T. \end{aligned}$$

We now show that the constraint of satisfying the Lagrange–d'Alembert principle leads to the following problem formulation, where the rigid body equations of motion replace the Lagrange–d'Alembert principle.

Problem 2.2. *Minimize*

$$(6) \qquad \mathcal{J} \;=\; \frac{1}{2} \int_0^T \ll \boldsymbol{\tau}, \boldsymbol{\tau} \gg_* \mathrm{d}t$$

subject to

(1) *the* dynamics

$$(7) \qquad \begin{aligned} \dot{\mathbf{R}} &= \mathbf{R}\boldsymbol{\Omega} \\ \dot{\mathbf{M}} &= \mathrm{ad}^*_{\boldsymbol{\Omega}}\mathbf{M} + \boldsymbol{\tau} = [\mathbf{M}, \boldsymbol{\Omega}] + \boldsymbol{\tau}, \end{aligned}$$

where $\mathbf{M} = \mathbf{J}(\boldsymbol{\Omega}) \in \mathfrak{so}^*(3)$ *is the momentum,*

(2) *and the* boundary conditions

$$(8) \qquad \begin{aligned} \mathbf{R}(0) &= \mathbf{R}_0, \; \boldsymbol{\Omega}(0) = \boldsymbol{\Omega}_0, \\ \mathbf{R}(T) &= \mathbf{R}_T, \; \boldsymbol{\Omega}(T) = \boldsymbol{\Omega}_T. \end{aligned}$$

In the above, ad^* is the dual of the adjoint representation, ad, of $\mathfrak{so}(3)$ and is given by $\mathrm{ad}^*_{\boldsymbol{\xi}}\boldsymbol{\eta} = -[\boldsymbol{\xi}, \boldsymbol{\eta}] \in \mathfrak{so}^*(3)$, for all $\boldsymbol{\xi} \in \mathfrak{so}(3)$ and $\boldsymbol{\eta} \in \mathfrak{so}^*(3)$. Recall that the bracket is defined by $[\boldsymbol{\xi}, \boldsymbol{\omega}] = \boldsymbol{\xi}\boldsymbol{\omega} - \boldsymbol{\omega}\boldsymbol{\xi}$.

2.2. The Lagrange–d'Alembert Principle and the Rigid Body Equations of Motion. In this section we derive the forced rigid body equations of motion (equations (7)) from the Lagrange–d'Alembert principle. We begin by appending the constraint $\dot{\mathbf{R}} = \mathbf{R}\boldsymbol{\Omega}$ to the Lagrangian

$$\begin{aligned} 0 \;=\; & \boldsymbol{\delta} \int_0^T \left(\frac{1}{2} \langle \mathbf{J}\left(\boldsymbol{\Omega}\right), \boldsymbol{\Omega} \rangle + \left\langle \boldsymbol{\Lambda}, \mathbf{R}^{-1}\dot{\mathbf{R}} - \boldsymbol{\Omega} \right\rangle \right) \mathrm{d}t \\ & + \int_0^T \langle \boldsymbol{\tau}, \mathbf{W} \rangle \, \mathrm{d}t, \end{aligned}$$

where $\boldsymbol{\Lambda} \in \mathfrak{so}^*(3)$ is a Lagrange multiplier. Taking variations, we obtain

$$\begin{aligned} 0 \;=\; & \int_0^T \left(\left\langle -\dot{\boldsymbol{\Lambda}} - [\boldsymbol{\Omega}, \boldsymbol{\Lambda}] + \boldsymbol{\tau}, \mathbf{W} \right\rangle + \langle -\boldsymbol{\Lambda} + \mathbf{J}\left(\boldsymbol{\Omega}\right), \boldsymbol{\delta}\boldsymbol{\Omega} \rangle \right) \mathrm{d}t \\ & + \left[\langle \boldsymbol{\Lambda}, \mathbf{W}(t) \rangle \right]_0^T. \end{aligned}$$

The term outside the integral vanishes by virtue of the boundary conditions on $\mathbf{W}(t)$. Since $\boldsymbol{\delta\Omega}$ and $\mathbf{W}$ are arbitrary and independent, we must have $\big\langle -\boldsymbol{\Lambda} + \mathbf{J}(\boldsymbol{\Omega}), \boldsymbol{\delta\Omega}\big\rangle = 0$ and hence $\boldsymbol{\Lambda} = \mathbf{J}(\boldsymbol{\Omega})$ which is the body angular momentum. For the remainder of this section, we set $\mathbf{M} = \boldsymbol{\Lambda} = \mathbf{J}(\boldsymbol{\Omega})$. The first term in the integrand gives us the forced second order dynamics of the rigid body which is

$$\dot{\mathbf{M}} = [\mathbf{M}, \boldsymbol{\Omega}] + \boldsymbol{\tau}.$$

This completes the proof that the Problem (2.1) is equivalent to Problem (2.2).

A Direct Approach. We now give a direct derivation that does not involve the use of Lagrange multipliers. This approach can be found in Section 13.5 in [21]. First, we take variations of the kinematic condition $\boldsymbol{\Omega} = \mathbf{R}^{-1}\dot{\mathbf{R}}$ to obtain $\boldsymbol{\delta\Omega} = -\mathbf{R}^{-1}(\boldsymbol{\delta}\mathbf{R})\mathbf{R}^{-1}\dot{\mathbf{R}} + \mathbf{R}^{-1}\left(\dot{\boldsymbol{\delta}\mathbf{R}}\right)$. As defined previously, we have $\mathbf{W} = \mathbf{R}^{-1}\boldsymbol{\delta}\mathbf{R}$ and, therefore, $\dot{\mathbf{W}} = -\mathbf{R}^{-1}\dot{\mathbf{R}}\mathbf{R}^{-1}\boldsymbol{\delta}\mathbf{R} + \mathbf{R}^{-1}\dot{\boldsymbol{\delta}\mathbf{R}} = -\boldsymbol{\Omega}\mathbf{W} + \mathbf{R}^{-1}\dot{\boldsymbol{\delta}\mathbf{R}}$, since $\boldsymbol{\delta}\dot{\mathbf{R}} = \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\delta}\mathbf{R}$ (see for example [22], page 52). Hence, we have

(9) $$\boldsymbol{\delta\Omega} = -\mathbf{W}\boldsymbol{\Omega} + \boldsymbol{\Omega}\mathbf{W} + \dot{\mathbf{W}} = \mathrm{ad}_{\Omega}\mathbf{W} + \dot{\mathbf{W}}.$$

Taking variations of the Lagrange–d'Alembert principle we obtain

$$\int_0^T \langle \mathbf{J}(\boldsymbol{\Omega}), \boldsymbol{\delta\Omega}\rangle + \langle \boldsymbol{\tau}, \mathbf{W}\rangle \, \mathrm{d}t = 0.$$

Using the variation in equation (9) and integrating by parts, we obtain

$$0 = \int_0^T \left\langle -\dot{\mathbf{M}} + \mathrm{ad}^*_{\Omega}\mathbf{M} + \boldsymbol{\tau}, \mathbf{W}\right\rangle \mathrm{d}t + \left[\langle \mathbf{J}(\boldsymbol{\Omega}), \mathbf{W}(t)\rangle\right]_0^T,$$

where we have used the identity

(10) $$\langle \boldsymbol{\eta}, \mathrm{ad}_{\boldsymbol{\omega}}\boldsymbol{\xi}\rangle = \langle \mathrm{ad}^*_{\boldsymbol{\omega}}\boldsymbol{\eta}, \boldsymbol{\xi}\rangle, \ \boldsymbol{\eta} \in \mathfrak{so}^*(3), \ \boldsymbol{\omega}, \boldsymbol{\xi} \in \mathfrak{so}(3).$$

This gives the desired result, with $\mathbf{M} = \mathbf{J}(\boldsymbol{\Omega})$.

In Section 2.3, we demonstrate how the necessary conditions for Problem (2.2) are derived using a variational approach.

2.3. Continuous-Time Variational Optimal Control Problem. A direct variational approach is used here to obtain the differential equation that satisfies the optimal control Problem (2.2).

A Second Order Direct Approach. "Second order" is used here to reflect the fact that we now study variations of second order dynamical equations as opposed to the kinematic direct approach studied in Section 2.2. We now give the resulting necessary conditions using a direct approach as in [21]. We already computed the variations of $\mathbf{R}$ and $\boldsymbol{\Omega}$. These were as follows: $\boldsymbol{\delta}\mathbf{R} = \mathbf{R}\mathbf{W}$ and $\boldsymbol{\delta\Omega} = \mathrm{ad}_{\Omega}\mathbf{W} + \dot{\mathbf{W}}$. We now compute the variation of $\dot{\mathbf{M}}$ with the goal of obtaining the proper variations for $\boldsymbol{\tau}$:

$$\boldsymbol{\delta}\dot{\mathbf{M}} \ = \ \mathbf{J}\left(\boldsymbol{\delta}\dot{\boldsymbol{\Omega}}\right) = \mathbf{J}\left(\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\delta\Omega} + \mathcal{R}(\mathbf{W}, \boldsymbol{\Omega})\boldsymbol{\Omega}\right),$$

where $\mathcal{R}$ is the curvature tensor on SO(3). The curvature tensor $\mathcal{R}$ arises due to the identity (see [22], page 52)

$$\frac{\partial}{\partial\epsilon}\frac{\partial}{\partial t}\mathbf{Y} - \frac{\partial}{\partial t}\frac{\partial}{\partial\epsilon}\mathbf{Y} = \mathcal{R}(\mathbf{W}, \mathbf{Y})\boldsymbol{\Omega},$$

where $\mathbf{Y} \in \mathsf{TSO}(3)$ is any vector field along the curve $\mathbf{R}(t) \in \mathsf{SO}(3)$. Taking variations of $\dot{\mathbf{M}} = \mathrm{ad}^*_{\mathbf{\Omega}}\mathbf{M} + \boldsymbol{\tau}$, we obtain $\delta\dot{\mathbf{M}} = \mathrm{ad}^*_{\boldsymbol{\delta\Omega}}\mathbf{M} + \mathrm{ad}^*_{\mathbf{\Omega}}\boldsymbol{\delta}\mathbf{M} + \boldsymbol{\delta\tau}$. We now have the desired variation in $\boldsymbol{\tau}$:

$$(11) \qquad \boldsymbol{\delta\tau} \;=\; \mathbf{J}\left(\mathcal{R}\left(\mathbf{W},\mathbf{\Omega}\right)\mathbf{\Omega}\right) + \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{J}\left(\boldsymbol{\delta\Omega}\right) - \mathrm{ad}^*_{\boldsymbol{\delta\Omega}}\mathbf{M} - \mathrm{ad}^*_{\mathbf{\Omega}}\boldsymbol{\delta}\mathbf{M}.$$

Taking variations of the cost functional (6) we obtain:

$$\boldsymbol{\delta}\mathcal{J} = \int_0^T \left(\left\langle \mathbf{J}(\ddot{\varsigma}) - \mathrm{ad}^*_{\mathbf{\Omega}}\left(\mathbf{J}(\dot{\varsigma})\right) + \dot{\boldsymbol{\eta}} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{ad}^*_{\boldsymbol{\varsigma}}\mathbf{M}\right) \right.\right.$$
$$\left.\left. + \left[\mathcal{R}\left(\mathbf{J}(\varsigma)^\sharp, \mathbf{\Omega}\right)\mathbf{\Omega}\right]^\flat + \mathrm{ad}^*_{\mathbf{\Omega}}\mathrm{ad}^*_{\boldsymbol{\varsigma}}\mathbf{M} - \mathrm{ad}^*_{\mathbf{\Omega}}\boldsymbol{\eta}, \mathbf{W}\right\rangle \right)\mathrm{d}t,$$

where $\varsigma = \boldsymbol{\tau}^\sharp \in \mathfrak{so}(3)$ and $\boldsymbol{\eta} = \mathbf{J}\left(\mathrm{ad}_{\mathbf{\Omega}}\varsigma\right) \in \mathfrak{so}^*(3)$. In obtaining the above expression, we have used integration by parts and the boundary conditions (8), equations (9) and (11), and the identities (1), (2) and (10). Hence, we have the following theorem.

Theorem 2.1. *The necessary optimality conditions for the problem of minimizing (6) subject to the dynamics (7) and the boundary conditions (8) are given by the single fourth order[2] differential equation*

$$0 \;=\; \mathbf{J}(\ddot{\varsigma}) - \mathrm{ad}^*_{\mathbf{\Omega}}\left(\mathbf{J}(\dot{\varsigma})\right) + \dot{\boldsymbol{\eta}} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{ad}^*_{\boldsymbol{\varsigma}}\mathbf{M}\right)$$
$$+ \left[\mathcal{R}\left((\mathbf{J}(\varsigma))^\sharp, \mathbf{\Omega}\right)\mathbf{\Omega}\right]^\flat + \mathrm{ad}^*_{\mathbf{\Omega}}\left(\mathrm{ad}^*_{\boldsymbol{\varsigma}}\mathbf{M}\right) - \mathrm{ad}^*_{\mathbf{\Omega}}\boldsymbol{\eta},$$

as well as the equations (7) and the boundary conditions (8), where ς and $\boldsymbol{\eta}$ are as defined above.

Note that for a compact semi-simple Lie group G with Lie algebra $\mathfrak{g}$, the curvature tensor, with respect to a bi-invariant metric, is given by (see [22]):

$$(12) \qquad\qquad \mathcal{R}\left(\mathbf{X},\mathbf{Y}\right)\mathbf{Z} = \frac{1}{4}\mathrm{ad}_{\mathrm{ad}_{\mathbf{X}}\mathbf{Y}}\mathbf{Z},$$

for all $\mathbf{X},\mathbf{Y},\mathbf{Z} \in \mathfrak{g}$.

Using a Lagrange multiplier approach as in Section 2.2, we may show that the result of Theorem 2.1 is equivalent to the following theorem.

Theorem 2.2. *The necessary optimality conditions for the problem of minimizing (6) subject to the dynamics (7) and the boundary conditions (8) are given by*

$$(13) \qquad \begin{aligned} \boldsymbol{\tau} &= \mathbf{\Lambda}_2 \\ \dot{\mathbf{\Lambda}}_1 &= \left[\mathcal{R}\left(\mathbf{J}\left(\mathbf{\Lambda}_2\right)^\sharp, \mathbf{\Omega}\right)\mathbf{\Omega}\right]^\flat + \mathrm{ad}^*_{\mathbf{\Omega}}\mathbf{\Lambda}_1 \\ \dot{\mathbf{\Lambda}}_2 &= -\mathbf{J}^{-1}\left(\mathbf{\Lambda}_1\right) - \mathrm{ad}_{\mathbf{\Omega}}\mathbf{\Lambda}_2 + \mathbf{J}^{-1}\left(\mathrm{ad}^*_{\mathbf{\Lambda}_2}\mathbf{M}\right) \end{aligned}$$

as well as the equations (7) and the boundary conditions (8), where the Lagrange multipliers $\mathbf{\Lambda}_1 \in \mathfrak{so}^*(3)$, $\mathbf{\Lambda}_2 \in \mathfrak{so}(3)$ correspond to the kinematic and dynamics constraints (7), respectively.

---

[2] Second order in $\boldsymbol{\tau}$ and fourth order in $\mathbf{R}$.

Remark 2.1. *Note that the equations of motion that arise from the Lagrange–d'Alembert principle are used to define the dynamic constraints. So, in effect, we are minimizing $\mathcal{J}$ subject to satisfying the Lagrange–d'Alembert principle for the rigid body. Analogously, the discrete version of the Lagrange–d'Alembert principle will be used to derive the discrete equations of motion in the discrete optimal control problem to be studied in Section 3.3. This view is in line with the approach in [11] in that we do not discretize the equations of motion directly, but, instead, we discretize the Lagrange–d'Alembert principle. These two approaches are not equivalent in general.*

Corollary 2.1. *The necessary optimality conditions of Theorem 2.1 are equivalent to the necessary conditions of Theorem 2.2.*

Proof. In Theorem 2.2, differentiate $\mathbf{\Lambda}_2$ once and then use all three differential equations to replace $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$ with expressions involving only $\boldsymbol{\tau}$, $\mathbf{M}$ and $\boldsymbol{\Omega}$.   □

3. Discrete-Time Results

3.1. **Problem Formulation.** In this section we give the discrete version of the problem introduced in Section 2.1. So, we consider minimizing the norm squared of the control torque $\boldsymbol{\tau}$ subject to satisfaction of the discrete Lagrange–d'Alembert principle for the rigid body whose configuration and body angular velocity at time step t_k are given by $\mathbf{R}_k \in \mathsf{SO}(3)$ and $\boldsymbol{\Omega}_k \in \mathfrak{so}(3)$, respectively. The kinematic constraint may be expressed as

$$(14) \qquad \mathbf{R}_{k+1} = \mathbf{R}_k \exp\left(h\boldsymbol{\Omega}_k\right) = \mathbf{R}_k \mathbf{g}_k,$$

where h is the integration time step, $\exp : \mathfrak{so}(3) \to \mathsf{SO}(3)$ is the exponential map and $\mathbf{g}_k = \exp(h\boldsymbol{\Omega}_k)$. The boundary conditions are given by $(\mathbf{R}_0^*, \boldsymbol{\Omega}_0^*)$ and $(\mathbf{R}_N^*, \boldsymbol{\Omega}_{N-1}^*)$, where $t_0 = 0$ and $N = T/h$ is such that $t_N = T$.

More generally, one considers the ansatz $\mathbf{R}_{k+1} = \mathbf{R}_k \exp\left(\boldsymbol{\Omega}(h)\right)$, where $\boldsymbol{\Omega}(\cdot)$ is an interpolatory curve in $\mathfrak{so}(3)$ parameterized by the angular velocity at internal nodal points. This allows one to construct Lie group variational integrators of arbitrarily high order [17]. To simplify the subsequent treatment, we adopt (14) as the kinematic constraint, which yields a first-order accurate Lie symplectic Euler method, which will nevertheless have effective order two as it is symplectically conjugate to the second-order accurate Lie Störmer–Verlet method (see, §4).

The reason we constrain $\boldsymbol{\Omega}$ at $t = h(N-1)$ instead of at $t = hN$ will become clear when we derive the discrete equations of motion in Section 3.2. A simple explanation for this is that a constraint on $\boldsymbol{\Omega}_k \in \mathfrak{so}(3)$ corresponds, by left translations to a constraint on $\dot{\mathbf{R}}_k \in \mathsf{T}_{\mathbf{R}_k}\mathsf{SO}(3)$. In turn, in the discrete setting and depending on the choice of discretization, this corresponds to a constraint on the neighboring discrete points $\ldots, \mathbf{R}_{k-2}, \mathbf{R}_{k-1}, \mathbf{R}_{k+1}, \mathbf{R}_{k+2}, \ldots$. With our choice of discretization (equation (14)), this corresponds to constraints on $\mathbf{R}_k$ and $\mathbf{R}_{k+1}$. Hence, to ensure that the effect of the terminal constraint on $\boldsymbol{\Omega}$ is correctly accounted for, the constraint must be imposed on $\boldsymbol{\Omega}_{N-1}$, which entails some constraints on variations at both $\mathbf{R}_{N-1}$ and $\mathbf{R}_N$. We will return to this point later in the paper.

The discrete kinematic constraint ensures that the sequence $\mathbf{R}_k$ stays on the rotation group, since the exponential of the angular velocity matrix $\boldsymbol{\Omega}_k$, which is in the algebra $\mathfrak{so}(3)$, is a rotation matrix, and the rotation group is closed under matrix multiplication. This is natural to do in the context of discrete variational numerical solvers (for both initial value and two point boundary value problems).

Following the methodology of [11], we have the following optimal control problem.

Problem 3.1. *Minimize*

$$(15) \qquad \mathcal{J} \;\; = \;\; \sum_{k=0}^{N} \frac{1}{2} \ll \boldsymbol{\tau}_k, \boldsymbol{\tau}_k \gg_*$$

subject to

(1) *satisfying the* discrete Lagrange–d'Alembert principle*:*

$$(16) \qquad \boldsymbol{\delta} \sum_{k=0}^{N-1} \frac{1}{2} \left\langle \mathbf{J}\left(\boldsymbol{\Omega}_k\right), \boldsymbol{\Omega}_k \right\rangle + \sum_{k=0}^{N} \left\langle \boldsymbol{\tau}_k, \mathbf{W}_k \right\rangle = 0,$$

subject to $\mathbf{R}_0 = \mathbf{R}_0^*$, $\mathbf{R}_N = \mathbf{R}_N^*$ and $\mathbf{R}_{k+1} = \mathbf{R}_k \mathbf{g}_k$, $k = 0, 1, \ldots, N-1$, where $\mathbf{W}_k$ is the variation vector field at time step t_k satisfying $\boldsymbol{\delta} \mathbf{R}_k = \mathbf{R}_k \mathbf{W}_k$,

(2) *and the* boundary conditions

$$(17) \qquad \begin{aligned} \mathbf{R}_0 &= \mathbf{R}_0^*, \ \boldsymbol{\Omega}_0 = \boldsymbol{\Omega}_0^*, \\ \mathbf{R}_N &= \mathbf{R}_N^*, \ \boldsymbol{\Omega}_{N-1} = \boldsymbol{\Omega}_{N-1}^*. \end{aligned}$$

In Problem 3.1, the discrete Lagrange–d'Alembert principle is used to derive the equations of motion for the rigid body with initial and terminal configuration constraints. Hence, we get a two point boundary value problem. The full configuration and velocity boundary conditions come into the picture when we study the optimal control problem. We will show that the constraint of satisfying the Lagrange–d'Alembert principle in Problem 3.1 leads to the following problem formulation, where the discrete rigid body equations of motion replace the Lagrange–d'Alembert principle constraint. Only when addressing the following optimal control problem will we need to include the velocity boundary conditions in the derivation.

Problem 3.2. *Minimize*

$$(18) \qquad \mathcal{J} \;\; = \;\; \sum_{k=0}^{N} \frac{1}{2} \ll \boldsymbol{\tau}_k, \boldsymbol{\tau}_k \gg_*$$

subject to

(1) *the* discrete dynamics

$$\mathbf{R}_{k+1} = \mathbf{R}_k \mathbf{g}_k, \ k = 0, \ldots, N-1$$
$$\mathbf{M}_k = \mathrm{Ad}_{\mathbf{g}_k}^* \left(h \boldsymbol{\tau}_k + \mathbf{M}_{k-1} \right), \ k = 1, \ldots, N-1,$$
$$(19) \qquad \mathbf{M}_k = \mathbf{J}\left(\boldsymbol{\Omega}_k\right), \ k = 0, \ldots, N-1,$$

(2) *and the* boundary conditions

$$(20) \qquad \begin{aligned} \mathbf{R}_0 &= \mathbf{R}_0^*, \ \boldsymbol{\Omega}_0 = \boldsymbol{\Omega}_0^*, \\ \mathbf{R}_N &= \mathbf{R}_N^*, \ \boldsymbol{\Omega}_{N-1} = \boldsymbol{\Omega}_{N-1}^*. \end{aligned}$$

Regarding terminal velocity conditions, note that in the second of equations (19) if we let $k = N$ we find that $\boldsymbol{\Omega}_N$ appears in the equation. A constraint on $\boldsymbol{\Omega}_N$ dictates constraints at the points $\mathbf{R}_N$ and $\mathbf{R}_{N+1}$ through the first equation in (19). Since we only consider time points up to $t = Nh$, we can not allow $k = N$ in the second of equations (19) and hence our terminal velocity constraints are posed in terms of $\boldsymbol{\Omega}_{N-1}$ instead of $\boldsymbol{\Omega}_N$.

As mentioned above, $\mathbf{W}_k$ is a variation vector field associated with the perturbed group element $\mathbf{R}_k^\epsilon$. Likewise, we need to define a variation vector field associated with the element $\mathbf{g}_k = \exp(h\mathbf{\Omega}_k)$. First, let the perturbed variable $\mathbf{g}_k^\epsilon$ be defined by

$$(21) \qquad \mathbf{g}_k^\epsilon = \mathbf{g}_k \exp(\epsilon h \boldsymbol{\delta\Omega}_k),$$

where

$$\boldsymbol{\delta\Omega}_k = \left. \frac{\partial \mathbf{\Omega}_k^\epsilon}{\partial \epsilon} \right|_{\epsilon=0}.$$

Note that $\mathbf{g}_k^\epsilon|_{\epsilon=0} = \mathbf{g}_k$ as desired. Moreover, we have

$$(22) \qquad \boldsymbol{\delta}\mathbf{g}_k = \mathbf{g}_k (h\boldsymbol{\delta\Omega}_k) \exp(\epsilon h \boldsymbol{\delta\Omega}_k)|_{\epsilon=0} = h\mathbf{g}_k \boldsymbol{\delta\Omega}_k.$$

This will be needed later when taking variations.

3.2. **The Discrete Lagrange–d'Alembert Principle and the Rigid Body Equations of Motion.** In this section we derive the discrete forced rigid body equations of motion (equations (19)) starting with the discrete Lagrange–d'Alembert principle. We begin by rewriting the kinematic constraint as $\exp^{-1}\left(\mathbf{R}_k^{-1}\mathbf{R}_{k+1}\right) = h\mathbf{\Omega}_k$, which is an expression on the Lie algebra $\mathfrak{so}(3)$, and appending it to the Lagrangian

$$
\begin{aligned}
0 \;=\; & \sum_{k=0}^{N} \langle \boldsymbol{\tau}_k, \mathbf{W}_k \rangle + \delta \sum_{k=0}^{N-1} \left(\frac{1}{2} \langle \mathbf{J}\left(\mathbf{\Omega}_k\right), \mathbf{\Omega}_k \rangle \right. \\
& \left. + \left\langle \mathbf{M}_k, \frac{1}{h} \exp^{-1}\left(\mathbf{R}_k^{-1}\mathbf{R}_{k+1}\right) - \mathbf{\Omega}_k \right\rangle \right),
\end{aligned}
$$

where $\mathbf{M}_k \in \mathfrak{so}^*(3)$, $k = 0, 1, \dots, N-1$, are Lagrange multipliers. These multipliers enforce the kinematic discrete equations. We could have added additional terms to enforce the configuration boundary conditions, allowing for $\mathbf{W}_0$ and $\mathbf{W}_N$ to be arbitrary and non-zero. Instead, we elect to enforce the constraints by requiring $\mathbf{W}_0 = 0$ and $\boldsymbol{\delta\Omega}_0 = 0$. These two approaches are of course equivalent.

Taking variations of $\frac{1}{h} \exp^{-1}\left(\mathbf{R}_k^{-1}\mathbf{R}_{k+1}\right) - \mathbf{\Omega}_k = 0$ is equivalent to taking variations of the original expression $\mathbf{R}_k^{-1}\mathbf{R}_{k+1} = \exp\left(h\mathbf{\Omega}_k\right)$, and is easier to compute since it as an expression over the Lie algebra. Once the variations are computed, one can easily obtain the Lie algebra-equivalent of the variations as follows. First take variations of the kinematics (14) to get $-\mathbf{R}_k^{-1}\left(\boldsymbol{\delta}\mathbf{R}_k\right)\mathbf{R}_k^{-1}\mathbf{R}_{k+1} + \mathbf{R}_k^{-1}\boldsymbol{\delta}\mathbf{R}_{k+1} = h\mathbf{g}_k \cdot \boldsymbol{\delta\Omega}_k$, which is equivalent to $-\mathbf{W}_k\mathbf{g}_k + \mathbf{g}_k\mathbf{W}_{k+1} = h\mathbf{g}_k\boldsymbol{\delta\Omega}_k$, or

$$(23) \qquad \boldsymbol{\delta\Omega}_k = \frac{1}{h}\left[-\mathrm{Ad}_{\mathbf{g}_k^{-1}}\mathbf{W}_k + \mathbf{W}_{k+1} \right].$$

Note that this is an expression over the Lie algebra $\mathfrak{so}(3)$.

After simple algebraic and re-indexing operations, the Lagrange–d'Alembert principle gives

$$0 = \left\langle \boldsymbol{\tau}_0 - \frac{1}{h} \mathrm{Ad}^*_{\mathbf{g}_0^{-1}} \mathbf{M}_0, \mathbf{W}_0 \right\rangle + \left\langle \boldsymbol{\tau}_N + \frac{1}{h} \mathbf{M}_{N-1}, \mathbf{W}_N \right\rangle$$
$$+ \sum_{k=0}^{N-1} \langle \mathbf{J}(\boldsymbol{\Omega}_k) - \mathbf{M}_k, \boldsymbol{\delta\Omega}_k \rangle$$
$$+ \sum_{k=1}^{N-1} \left\langle \boldsymbol{\tau}_k - \frac{1}{h} \mathrm{Ad}^*_{\mathbf{g}_k^{-1}} \mathbf{M}_k + \frac{1}{h} \mathbf{M}_{k-1}, \mathbf{W}_k \right\rangle.$$

By the boundary conditions $\mathbf{R}_0 = \mathbf{R}_0^*$ and $\mathbf{R}_N = \mathbf{R}_N^*$, we have $\mathbf{W}_0 = 0$ and $\mathbf{W}_N = 0$. Since $\boldsymbol{\delta\Omega}_k$, $k = 0, \ldots, N-1$, and $\mathbf{W}_k$, $k = 1, \ldots, N-1$, are arbitrary and independent, then the Lagrange–d'Alembert principle requires that the equations (19) hold true. The variables $\mathbf{M}_k$, $k = 0, \ldots, N-1$, are of course nothing but the discrete angular momentum of the rigid body. The equations (19) can be viewed in two ways. The first is to consider the two point boundary value problem where we retain the terminal condition on $\mathbf{R}_N$. In this case a (constrained) variety of a combination of control torques $\boldsymbol{\tau}_k$, $k = 0, \ldots, N$, and initial velocity conditions $\boldsymbol{\Omega}_0$ can be chosen to drive the rigid body from the initial condition $\mathbf{R}_0$ to the terminal condition $\mathbf{R}_N$. The second view is to treat it as an initial value problem by ignoring any terminal configuration constraints. In this case $\mathbf{W}_N \neq 0$ and any combination of control torques $\boldsymbol{\tau}_k$, $k = 0, \ldots, N$, and initial velocity conditions $\boldsymbol{\Omega}_0$ can be chosen freely.

A Direct Approach. Taking direct variations of the cost functional we obtain

$$0 = \left\langle \boldsymbol{\tau}_0 - \frac{1}{h} \mathrm{Ad}^*_{\mathbf{g}_0^{-1}} \mathbf{J}(\boldsymbol{\Omega}_0), \mathbf{W}_0 \right\rangle$$
$$+ \left\langle \boldsymbol{\tau}_N + \frac{1}{h} \mathbf{J}(\Omega_{N-1}), \mathbf{W}_N \right\rangle$$
$$+ \sum_{k=1}^{N-1} \left\langle \boldsymbol{\tau}_k - \frac{1}{h} \mathrm{Ad}^*_{\mathbf{g}_k^{-1}} \mathbf{J}(\boldsymbol{\Omega}_k) + \frac{1}{h} \mathbf{J}(\boldsymbol{\Omega}_{k-1}), \mathbf{W}_k \right\rangle.$$

where we have used equation (23). This gives the same equations of motion as those in equation (19).

Simulation Results. To test our results, we re-write the discrete equations (19) for the subgroup $\mathsf{SO}(2)$. For $\mathsf{SO}(2)$ we have

$$(24) \qquad \mathbf{R}_k = \begin{bmatrix} \cos\theta_k & -\sin\theta_k \\ \sin\theta_k & \cos\theta_k \end{bmatrix}, \ \boldsymbol{\Omega}_k = \begin{bmatrix} 0 & -\omega_k \\ \omega_k & 0 \end{bmatrix}$$

and

$$(25) \qquad \exp(\boldsymbol{\Omega}_k) = \begin{bmatrix} \cos\omega_k & -\sin\omega_k \\ \sin\omega_k & \cos\omega_k \end{bmatrix}.$$

The inertia operation is simply given by

$$(26) \qquad \mathbf{J}(\boldsymbol{\Omega}_k) = \begin{bmatrix} 0 & -I\omega_k \\ I\omega_k & 0 \end{bmatrix},$$

where I is the mass moment of inertia of the body about the out-of-plane axis. One can check that $\mathrm{Ad}_{\exp(\boldsymbol{\omega})}\boldsymbol{\xi} = \boldsymbol{\xi}$ and that $\mathrm{Ad}^*_{\exp(\boldsymbol{\omega})}\boldsymbol{\eta} = \boldsymbol{\eta}$, for all $\boldsymbol{\xi}$, $\boldsymbol{\omega} \in \mathfrak{so}(2)$ and $\boldsymbol{\eta} \in \mathfrak{so}^*(2)$.

Then the equations (19) (treated as an initial value problem) are given for $\mathsf{SO}(2)$ by

$$
\begin{aligned}
\theta_{k+1} &= \theta_k + h\omega_k, \ k = 0, \dots, N-1 \\
\omega_k &= \frac{h}{I}\tau_k + \omega_{k-1}, \ k = 1, \dots, N-1
\end{aligned}
$$
(27)

in addition to the initial conditions $\theta_0 = \theta_0^*$, $\omega_0 = \omega_0^*$.

To verify the accuracy of our numerical computation, we give the corresponding continuous-time equations of motion for the planar rigid body on $\mathsf{SO}(2)$ using equations (7). The Lie bracket on $\mathsf{SO}(2)$ is identically equal to zero. Hence, one can check that the equations (7) are given by $\dot{\theta} = \omega$, $\dot{\omega} = \frac{\tau}{I}$, where θ, ω and τ are the continuous time angular position, velocity and torque, respectively. We integrate the equations using the torque $\tau(t) = \sin\left(\frac{\pi t}{2}\right)$, $t \in [0, T]$. We use the following parameters for our simulations: $T = 10$, $I = 1$, $\theta(0) = 3$, $\omega(0) = 4$ and we try three different time steps corresponding to $N = 1000, 1500, 2000$. The error between the continuous- and discrete-time values of θ and ω are given in Figure (1). Note that the accuracy of the simulation improves with increasing N.

Remark 3.1. *Note that the discrete-time equations (27) correspond to the Euler approximation for the equations of motion. This is a check that our method returns something familiar for a simple example as the planar rigid body. However, we emphasize that on $\mathsf{SO}(3)$ the discretization will not necessarily be equivalent to any of the classical discretization schemes. The discretization will generally result in a set of nonlinear implicit algebraic equations.*

3.3. **Discrete-Time Variational Optimal Control Problem.** We now address Problem 3.2 by first forming the appended cost functional:

$$
\begin{aligned}
\mathcal{J} &= \sum_{k=0}^{N} \frac{1}{2} \ll \boldsymbol{\tau}_k, \boldsymbol{\tau}_k \gg_* \\
&+ \sum_{k=0}^{N-1} \left\langle \boldsymbol{\Lambda}_k^1, -h\boldsymbol{\Omega}_k + \exp^{-1}\left(\mathbf{R}_k^{-1}\mathbf{R}_{k+1}\right) \right\rangle \\
&+ \sum_{k=1}^{N-1} \left\langle \mathbf{M}_k - \mathrm{Ad}^*_{\mathbf{g}_k}\left(h\boldsymbol{\tau}_k + \mathbf{M}_{k-1}\right), \boldsymbol{\Lambda}_k^2 \right\rangle.
\end{aligned}
$$

Writing $\mathbf{M}_k = \mathrm{Ad}^*_{\mathbf{g}_k}\left(h\boldsymbol{\tau}_k + \mathbf{M}_{k-1}\right)$ as $\mathbf{M}_k = \mathbf{g}_k^{-1}\left(h\boldsymbol{\tau}_k + \mathbf{M}_{k-1}\right)\mathbf{g}_k$ and taking variations of this expression, we obtain

$$
\begin{aligned}
\delta\mathbf{M}_k &= -h\delta\boldsymbol{\Omega}_k\mathbf{g}_k^{-1}\left(h\boldsymbol{\tau}_k + \mathbf{M}_{k-1}\right)\mathbf{g}_k \\
&\quad + \mathbf{g}_k^{-1}\left(h\delta\boldsymbol{\tau}_k + \delta\mathbf{M}_{k-1}\right)\mathbf{g}_k \\
&\quad + h\mathbf{g}_k^{-1}\left(h\boldsymbol{\tau}_k + \mathbf{M}_{k-1}\right)\mathbf{g}_k\delta\boldsymbol{\Omega}_k \\
&= -h\delta\boldsymbol{\Omega}_k\mathbf{M}_k + \mathrm{Ad}^*_{\mathbf{g}_k}\left(h\delta\boldsymbol{\tau}_k + \mathbf{J}\left(\delta\boldsymbol{\Omega}_{k-1}\right)\right) \\
&\quad + h\mathbf{M}_k\delta\boldsymbol{\Omega}_k \\
&= \mathrm{Ad}^*_{\mathbf{g}_k}\left(h\delta\boldsymbol{\tau}_k + \mathbf{J}\left(\delta\boldsymbol{\Omega}_{k-1}\right)\right) + h\left[\mathbf{M}_k, \delta\boldsymbol{\Omega}_k\right].
\end{aligned}
$$
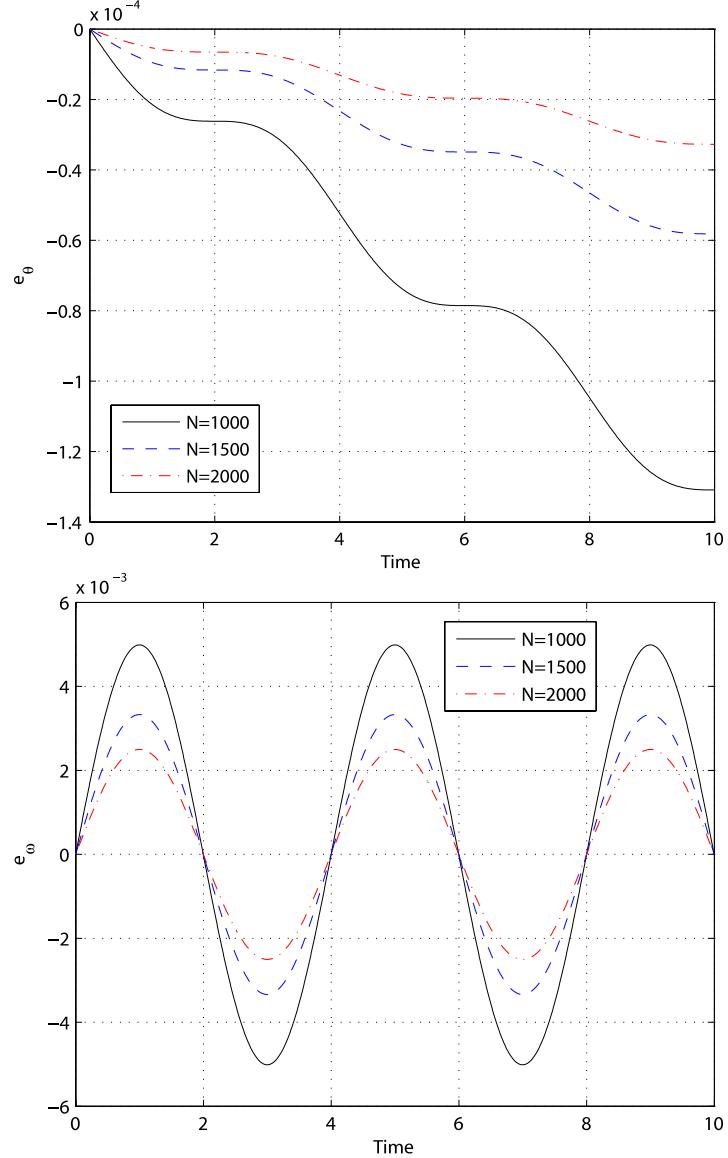
FIGURE 1. Error dynamics on $\mathsf{SO}(2)$.

In obtaining this expression we used the facts that $\boldsymbol{\delta}\mathbf{g}_k = h\mathbf{g}_k\boldsymbol{\delta}\boldsymbol{\Omega}_k$ and $\boldsymbol{\delta}\left(\mathbf{g}_k^{-1}\right) = -h\boldsymbol{\delta}\boldsymbol{\Omega}_k\mathbf{g}_k^{-1}$. The latter equality is obtained as follows. Taking variations of $\left(\mathbf{g}_k^{\epsilon}\right)^{-1}\left(\mathbf{g}_k^{\epsilon}\right) = \mathbf{I}$, we obtain

$$0 \;=\; \boldsymbol{\delta}\left(\mathbf{g}_k^{-1}\right)\mathbf{g}_k + h\mathbf{g}_k^{-1}\mathbf{g}_k\boldsymbol{\delta}\boldsymbol{\Omega}_k$$

which implies that

$$\boldsymbol{\delta}\left(\mathbf{g}_k^{-1}\right) = -h\boldsymbol{\delta}\boldsymbol{\Omega}_k\mathbf{g}_k^{-1},$$

which is the desired result.

We now also consider the velocity boundary conditions $\mathbf{\Omega}_0 = \mathbf{\Omega}_0^*$ and $\mathbf{\Omega}_{N-1} = \mathbf{\Omega}_{N-1}^*$. Note that variations in $\boldsymbol{\delta\Omega}_k$ directly induce variations in $\mathbf{W}_{k+1}$. In particular, if $k = 0$ and we have the initial constraints $\mathbf{R}_0 = \mathbf{R}_0^*$ and $\mathbf{\Omega}_0 = \mathbf{\Omega}_0^*$, then $\mathbf{W}_0 = 0$ and $\boldsymbol{\delta\Omega}_0 = 0$. Using these two equations in equation (23), we find that

$$(28) \qquad \mathbf{W}_1 = 0.$$

At $k = N, N-1$, the constraints $\mathbf{R}_N = \mathbf{R}_N^*$ and $\mathbf{\Omega}_{N-1} = \mathbf{\Omega}_{N-1}^*$ imply that $\mathbf{W}_N = 0$ and $\boldsymbol{\delta\Omega}_{N-1} = 0$. Using these two equations in equation (23), we find that

$$(29) \qquad \mathbf{W}_{N-1} = 0.$$

The observations stated in equations (28) and (29) are equivalent to having

$$(30) \qquad \mathbf{R}_1 = \mathbf{R}_0^* \exp\left(h\mathbf{\Omega}_0^*\right), \ \ \mathbf{R}_{N-1} = \mathbf{R}_N^* \exp\left(-h\mathbf{\Omega}_{N-1}^*\right).$$

Taking variations of the cost functional, we obtain

$$\begin{aligned}
\boldsymbol{\delta\mathcal{J}} = &\sum_{k=0}^{N} \left\langle \boldsymbol{\delta\tau}_k, \boldsymbol{\tau}_k^\sharp \right\rangle \\
&+ \sum_{k=0}^{N-1} \left\langle \mathbf{\Lambda}_k^1, -\boldsymbol{\delta\Omega}_k + \frac{1}{h} \left[-\mathrm{Ad}_{\mathbf{g}_k^{-1}} \mathbf{W}_k + \mathbf{W}_{k+1} \right] \right\rangle \\
&+ \sum_{k=1}^{N-1} \left\langle \mathbf{J}\left(\boldsymbol{\delta\Omega}_k\right) - \mathrm{Ad}_{\mathbf{g}_k}^* \left(h\boldsymbol{\delta\tau}_k + \mathbf{J}\left(\boldsymbol{\delta\Omega}_{k-1}\right)\right) \right. \\
&\qquad \left. -h\left[\mathbf{M}_k, \boldsymbol{\delta\Omega}_k\right], \mathbf{\Lambda}_k^2 \right\rangle,
\end{aligned}$$

where we have replaced $\boldsymbol{\delta M}_k$ with $\mathbf{J}\left(\boldsymbol{\delta\Omega}_k\right)$. Collecting terms, setting $\boldsymbol{\delta\mathcal{J}}$ to zero, and using the conditions $\mathbf{W}_0 = \mathbf{W}_1 = \mathbf{W}_{N-1} = \mathbf{W}_N = \boldsymbol{\delta\Omega}_0 = \boldsymbol{\delta\Omega}_{N-1} = 0$ and the fact that $\mathbf{J}(\cdot)$ is self-adjoint, we obtain the following theorem.

Theorem 3.1. *The necessary optimality conditions for the discrete Problem 3.2 are*

$$\begin{aligned}
\mathbf{R}_{k+1} &= \mathbf{R}_k \mathbf{g}_k, \ k = 1, \ldots, N-2 \\
\mathbf{M}_k &= \mathrm{Ad}_{\mathbf{g}_k}^* \left(h\boldsymbol{\tau}_k + \mathbf{M}_{k-1}\right), \ k = 1, \ldots, N-1 \\
0 &= \mathbf{\Lambda}_{k-1}^1 - \mathrm{Ad}_{\mathbf{g}_k^{-1}}^* \mathbf{\Lambda}_k^1, \ k = 2, \ldots, N-2 \\
(31) \qquad 0 &= -\mathbf{\Lambda}_k^1 + \mathbf{J}\left(\mathbf{\Lambda}_k^2\right) - \mathbf{J}\left(\mathrm{Ad}_{\mathbf{g}_{k+1}} \mathbf{\Lambda}_{k+1}^2\right) \\
&\quad +h\left[\mathbf{M}_k, \mathbf{\Lambda}_k^2\right], \ k = 1, \ldots, N-2 \\
\boldsymbol{\tau}_k &= h\left(\mathrm{Ad}_{\mathbf{g}_k} \mathbf{\Lambda}_k^2\right)^\flat, \ k = 1, \ldots, N-1 \\
\mathbf{M}_k &= \mathbf{J}\left(\mathbf{\Omega}_k\right), \ k = 0, \ldots, N-1,
\end{aligned}$$

and the boundary conditions

$$\begin{aligned}
\mathbf{R}_0 &= \mathbf{R}_0^*, \ \mathbf{R}_1 = \mathbf{R}_0^* \mathbf{g}_0^*, \ \mathbf{\Omega}_0 = \mathbf{\Omega}_0^* \\
\mathbf{R}_N &= \mathbf{R}_N^*, \ \mathbf{R}_{N-1} = \mathbf{R}_N^* \left(\mathbf{g}_{N-1}^*\right)^{-1}, \ \mathbf{\Omega}_{N-1} = \mathbf{\Omega}_{N-1}^* \\
\boldsymbol{\tau}_0 &= \boldsymbol{\tau}_N = 0,
\end{aligned}$$

where $\mathbf{g}_0^* = \exp(h\mathbf{\Omega}_0^*)$ and $\mathbf{g}_{N-1}^* = \exp\left(h\mathbf{\Omega}_{N-1}^*\right)$.

A Second Order Direct Approach. Analogous to the direct approach in continuous time, here we derive the necessary optimality conditions in a form that does not involve the use of Lagrange multipliers. Using equation (23) and taking variation of the second of equations (19), we obtain

$$
\begin{aligned}
\boldsymbol{\delta\tau}_k \;=\; & \mathrm{Ad}^*_{\mathbf{g}_k^{-1}}\left(\frac{1}{h^2}\mathbf{J}\left(\mathbf{W}_{k+1}-\mathrm{Ad}_{\mathbf{g}_k^{-1}}\mathbf{W}_k\right)\right. \\
& +\frac{1}{h}\left[\mathbf{W}_{k+1}-\mathrm{Ad}_{\mathbf{g}_k^{-1}}\mathbf{W}_k,\mathbf{J}\left(\boldsymbol{\Omega}_k\right)\right]\Big) \\
& -\frac{1}{h^2}\mathbf{J}\left(\mathbf{W}_k-\mathrm{Ad}_{\mathbf{g}_{k-1}^{-1}}\mathbf{W}_{k-1}\right),
\end{aligned}
$$

(32)

for $k=1,\ldots,N-1$. Taking variations of the cost functional (18) and substituting from equation (32) one obtains after a tedious but straight forward computation an expression for $\boldsymbol{\delta J}$ in terms of $\boldsymbol{\delta\tau}_k$:

$$
\begin{aligned}
\boldsymbol{\delta J} \;=\; & \sum_{k=1}^{N-1}\left[\left\langle\mathrm{Ad}^*_{\mathbf{g}_k^{-1}}\left(\frac{1}{h^2}\mathbf{J}\left(\mathbf{W}_{k+1}-\mathrm{Ad}_{\mathbf{g}_k^{-1}}\mathbf{W}_k\right)\right.\right.\right. \\
& \left.\left.+\frac{1}{h}\left[\mathbf{W}_{k+1}-\mathrm{Ad}_{\mathbf{g}_k^{-1}}\mathbf{W}_k,\mathbf{J}\left(\boldsymbol{\Omega}_k\right)\right]\right)-\frac{1}{h^2}\mathbf{J}\left(\mathbf{W}_k-\mathrm{Ad}_{\mathbf{g}_{k-1}^{-1}}\mathbf{W}_{k-1}\right),\boldsymbol{\tau}_k^\sharp\right\rangle\right] \\
& +\left\langle\boldsymbol{\delta\tau}_0,\boldsymbol{\tau}_0^\sharp\right\rangle+\left\langle\boldsymbol{\delta\tau}_N,\boldsymbol{\tau}_N^\sharp\right\rangle.
\end{aligned}
$$

When $\boldsymbol{\delta J}$ is equated to zero (and after some algebraic rearrangement), one can obtain the boundary conditions on $\boldsymbol{\tau}_0,\boldsymbol{\tau}_1,\boldsymbol{\tau}_{N-1},\boldsymbol{\tau}_N$ from the resulting equations below:

$$
\begin{aligned}
\boldsymbol{\tau}_0 \;=\;& 0 \\
0 \;=\;& -\frac{1}{h^2}\left(\mathbf{J}\left(\boldsymbol{\tau}_1^\sharp\right)+\mathrm{Ad}^*_{\mathbf{g}_1^{-1}}\mathbf{J}\left(\mathrm{Ad}_{\mathbf{g}_1^{-1}}\boldsymbol{\tau}_1^\sharp\right)\right) \\
& -\frac{1}{h}\mathrm{Ad}^*_{\mathbf{g}_1^{-1}}\left[\mathbf{J}\left(\boldsymbol{\Omega}_1\right),\mathrm{Ad}_{\mathbf{g}_1^{-1}}\left(\boldsymbol{\tau}_1^\sharp\right)\right] \\
0 \;=\;& -\frac{1}{h^2}\left(\mathbf{J}\left(\boldsymbol{\tau}_{N-1}^\sharp\right)+\mathrm{Ad}^*_{\mathbf{g}_{N-1}^{-1}}\mathbf{J}\left(\mathrm{Ad}_{\mathbf{g}_{N-1}^{-1}}\boldsymbol{\tau}_{N-1}^\sharp\right)\right) \\
& -\frac{1}{h}\mathrm{Ad}^*_{\mathbf{g}_{N-1}^{-1}}\left[\mathbf{J}\left(\boldsymbol{\Omega}_{N-1}\right),\mathrm{Ad}_{\mathbf{g}_{N-1}^{-1}}\left(\boldsymbol{\tau}_{N-1}^\sharp\right)\right] \\
\boldsymbol{\tau}_N \;=\;& 0
\end{aligned}
$$

as well as discrete evolution equations that are written in algebraic nonlinear form as:

$$
\begin{aligned}
0 \;=\;& -\frac{1}{h^2}\left(\mathbf{J}\left(\boldsymbol{\tau}_k^\sharp\right)-\mathrm{Ad}^*_{\mathbf{g}_k^{-1}}\mathbf{J}\left(\boldsymbol{\tau}_{k+1}^\sharp\right)\right. \\
& \left.-\mathbf{J}\left(\mathrm{Ad}_{\mathbf{g}_{k-1}^{-1}}\boldsymbol{\tau}_{k-1}^\sharp\right)+\mathrm{Ad}^*_{\mathbf{g}_k^{-1}}\mathbf{J}\left(\mathrm{Ad}_{\mathbf{g}_k^{-1}}\boldsymbol{\tau}_k^\sharp\right)\right) \\
& -\frac{1}{h}\left(\mathrm{Ad}^*_{\mathbf{g}_k^{-1}}\left[\mathbf{J}\left(\boldsymbol{\Omega}_k\right),\mathrm{Ad}_{\mathbf{g}_k^{-1}}\left(\boldsymbol{\tau}_k^\sharp\right)\right]\right. \\
& \left.-\frac{1}{h}\left[\mathbf{J}\left(\boldsymbol{\Omega}_{k-1}\right),\mathrm{Ad}_{\mathbf{g}_{k-1}^{-1}}\left(\boldsymbol{\tau}_{k-1}^\sharp\right)\right]\right),
\end{aligned}
$$

(33)

for $k=2,\ldots,N-2$.

The following section shows that while our discrete approximation (14) is formally first-order accurate, it is symplectically equivalent to the second-order accurate Störmer–Verlet method, and hence has effective order two.

4. Lie Symplectic Euler and Symplectic Equivalence

Notice that the discrete Lagrangian adopted in our paper is obtained by approximating the velocity as a constant over the timestep h, and by approximating the integral in time by $\int_{t_1}^{t_2} f(t)dt \approx (t_2 - t_1)f(t_1)$. In the Lie group setting, the constant angular velocity approximation corresponds to the condition,

$$\mathbf{R}_{k+1} = \mathbf{R}_k \exp(h\mathbf{\Omega}_k)$$

or equivalently,

$$\mathbf{\Omega}_k = \frac{1}{h}\exp^{-1}(\mathbf{R}_k^{-1}\mathbf{R}_{k+1}).$$

When we let $G = \mathbb{R}^n$, and we adopt the notation $(\mathbf{q}, \mathbf{v}) \in T\mathbb{R}^n$, we obtain,

$$\mathbf{v}_k = \frac{\mathbf{q}_{k+1} - \mathbf{q}_k}{h},$$

which is a usual finite-difference approximation for the velocity. Consider then a Lagrangian of the form,

$$L(\mathbf{q}, \mathbf{v}) = \frac{1}{2}\mathbf{v}^T M\mathbf{v} - V(\mathbf{q}).$$

Approximating the action integral from 0 to h using a constant velocity approximation and a quadrature formula, yields,

$$\int_0^h L(\mathbf{q}(t), \mathbf{v}(t))dt \approx \int_0^h L\Big(\mathbf{q}(t), \frac{\mathbf{q}_{k+1} - \mathbf{q}_k}{h}\Big)dt \approx hL\Big(\mathbf{q}_k, \frac{\mathbf{q}_{k+1} - \mathbf{q}_k}{h}\Big).$$

We then choose as our discrete Lagrangian,

$$L_d(\mathbf{q}_k, \mathbf{q}_{k+1}) = hL\Big(\mathbf{q}_k, \frac{\mathbf{q}_{k+1} - \mathbf{q}_k}{h}\Big) = h\left[\frac{1}{2}\Big(\frac{\mathbf{q}_{k+1} - \mathbf{q}_k}{h}\Big)^T M\Big(\frac{\mathbf{q}_{k+1} - \mathbf{q}_k}{h}\Big) - V(\mathbf{q}_k)\right].$$

The discrete Euler–Lagrange equations,

$$D_2 L_d(\mathbf{q}_{k-1}, \mathbf{q}_k) + D_1 L_d(\mathbf{q}_k, \mathbf{q}_{k+1}) = 0,$$

yields,

$$M\Big(\frac{\mathbf{q}_k - \mathbf{q}_{k-1}}{h}\Big) - M\Big(\frac{\mathbf{q}_{k+1} - \mathbf{q}_k}{h}\Big) - h\frac{\partial V}{\partial \mathbf{q}}(\mathbf{q}_k) = 0,$$

which induces an implicit update map $(\mathbf{q}_{k-1}, \mathbf{q}_k) \mapsto (\mathbf{q}_k, \mathbf{q}_{k+1})$. To obtain the corresponding Hamiltonian update map, we push-forward this algorithm to T^*Q by using the discrete fiber derivative $\mathbb{F}L_d : Q \times Q \to T^*Q$, which takes $(\mathbf{q}_k, \mathbf{q}_{k+1}) \mapsto (\mathbf{q}_{k+1}, D_2 L_d(\mathbf{q}_k, \mathbf{q}_{k+1}))$. In particular, we have that,

$$\mathbf{p}_{k+1} = D_2 L_d(\mathbf{q}_k, \mathbf{q}_{k+1}) = M\Big(\frac{\mathbf{q}_{k+1} - \mathbf{q}_k}{h}\Big),$$

which implies

(34) $$\mathbf{q}_{k+1} = \mathbf{q}_k + hM^{-1}\mathbf{p}_{k+1}.$$

This allows us to rewrite the discrete Euler–Lagrange equations as,

$$\mathbf{p}_k - \mathbf{p}_{k+1} - h\frac{\partial V}{\partial q}(\mathbf{q}_k) = 0,$$

or equivalently,

$$(35) \qquad \mathbf{p}_{k+1} = \mathbf{p}_k - h\frac{\partial V}{\partial q}(\mathbf{q}_k).$$

Now, (34) and (35) are precisely the symplectic Euler method applied to the corresponding Hamiltonian vector field, as we shall see.

The corresponding Hamiltonian is given by,

$$H(\mathbf{q}, \mathbf{p}) = \frac{1}{2}\mathbf{p}^T M^{-1}\mathbf{p} + V(\mathbf{q}).$$

Hamilton's equations yield,

$$\begin{pmatrix} \dot{\mathbf{q}} \\ \dot{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} \frac{\partial H}{\partial \mathbf{p}} \\ -\frac{\partial H}{\partial \mathbf{q}} \end{pmatrix} = \begin{pmatrix} M^{-1}\mathbf{p} \\ -\frac{\partial V}{\partial \mathbf{q}} \end{pmatrix}.$$

The symplectic Euler method has the form,

$$\mathbf{q}_{k+1} = \mathbf{q}_k + h\dot{\mathbf{q}}(\mathbf{q}_k, \mathbf{p}_{k+1}),$$
$$\mathbf{p}_{k+1} = \mathbf{p}_k + h\dot{\mathbf{p}}(\mathbf{q}_k, \mathbf{p}_{k+1}),$$

which yields,

$$\mathbf{q}_{k+1} = \mathbf{q}_k + hM^{-1}\mathbf{p}_{k+1},$$
$$\mathbf{p}_{k+1} = \mathbf{p}_k + h\left(-\frac{\partial V}{\partial q}(\mathbf{q}_k)\right),$$

which is precisely what we obtained in (34) and (35). This demonstrates that our method is the generalization of the symplectic Euler method to Lie groups, which has important numerical consequences. While symplectic Euler is formally first-order accurate, it is symplectically equivalent [27, 18] to the second-order accurate Störmer–Verlet method [6]. This means that one can obtain the Störmer–Verlet method F_{SV} by conjugating the symplectic Euler method F_{E} with a symplectic transformation T,

$$F_{\mathrm{SV}} = TF_{\mathrm{E}}T^{-1}.$$

In particular, numerical trajectories of symplectic Euler will shadow numerical trajectories obtained using Störmer–Verlet. Consider the implications of this symplectic equivalence for our discrete optimal control problem. Let the boundary conditions be specified by $\mathbf{q}_0, \mathbf{q}_N$, and assume that we use Störmer–Verlet to propagate the solution, then the boundary condition is expressed as, $\mathbf{q}_N = F_{\mathrm{SV}}^N \mathbf{q}_0 = (TF_{\mathrm{E}}T^{-1})^N \mathbf{q}_0 = TF_{\mathrm{E}}^N T^{-1}\mathbf{q}_0$, which is equivalent to $\tilde{q}_N = T^{-1}\mathbf{q}_N = F_{\mathrm{E}}^N T^{-1}\mathbf{q}_0 = F_{\mathrm{E}}^N \tilde{q}_0$. This implies that if we preprocess the boundary conditions $\mathbf{q}_0, \mathbf{q}_N$, to obtain $\tilde{q}_0 = T^{-1}\mathbf{q}_0, \tilde{q}_N = T^{-1}\mathbf{q}_N$, we could use symplectic Euler at the internal stages to propagate the states and costates, and then postprocess them to obtain the trajectory one would have obtained by using Störmer–Verlet.

In practice, the shadowing result imparts the symplectic Euler method with the same desirable qualitative properties as Störmer–Verlet, and it is not necessary to postprocess the numerical solutions in order to achieve accurate results. Since on an appropriate choice of charts, our Lie symplectic Euler method reduces to symplectic Euler in coordinates, it follows that there is a corresponding second-order Lie Störmer–Verlet method that our method is symplectically equivalent to, and in particular, our method has effective order two.

5. Numerical Approach and Results

The multiplier-free version of the first-order optimality equations, equation (33), in combination with the boundary conditions,

$$\mathbf{R}_0 = \mathbf{R}_0^*, \;\; \mathbf{R}_N = \mathbf{R}_N^*, \;\; \mathbf{\Omega}_0 = \mathbf{\Omega}_0^*, \;\; \text{and } \mathbf{\Omega}_{N-1} = \mathbf{\Omega}_{N-1}^*,$$

leave the torques $\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_{N-1}$, and the angular velocities $\mathbf{\Omega}_1, \ldots, \mathbf{\Omega}_{N-2}$ as unknowns. By substituting the relations $\mathbf{g}_k = \exp(h\mathbf{\Omega}_k)$, $\mathbf{M}_k = \mathbf{J}(\mathbf{\Omega}_k)$, we can rewrite the necessary conditions (33) as follows,

$$
\begin{aligned}
0 = & -\frac{1}{h^2}\Big(\mathbf{J}(\boldsymbol{\tau}_k^\sharp) - \mathrm{Ad}^*_{\exp(-h\mathbf{\Omega}_k)} \mathbf{J}(\boldsymbol{\tau}_{k+1}^\sharp) \\
& \quad - \mathbf{J}(\mathrm{Ad}_{\exp(-h\mathbf{\Omega}_{k-1})} \boldsymbol{\tau}_{k-1}^\sharp \\
& \quad + \mathrm{Ad}^*_{\exp(-h\mathbf{\Omega}_k)} \mathbf{J}(\mathrm{Ad}_{\exp(-h\mathbf{\Omega}_k)} \boldsymbol{\tau}_k^\sharp)) \Big) \\
& -\frac{1}{h}\Big(\mathrm{Ad}^*_{\exp(-h\mathbf{\Omega}_k)} \big[\mathbf{J}(\mathbf{\Omega}_k), \mathrm{Ad}_{\exp(-h\mathbf{\Omega}_k)}(\boldsymbol{\tau}_k^\sharp) \big] \\
& \quad - \frac{1}{h} \big[\mathbf{J}(\mathbf{\Omega}_{k-1}), \mathrm{Ad}_{\exp(-h\mathbf{\Omega}_{k-1})}(\boldsymbol{\tau}_{k-1}^\sharp) \big] \Big),
\end{aligned}
$$

where $k = 2, \ldots, N-2$, and the discrete evolution equations, given by line 2 of (31), can be written as

$$0 = \mathbf{J}(\mathbf{\Omega}_k) - \mathrm{Ad}^*_{\exp(h\mathbf{\Omega}_k)}(h\boldsymbol{\tau}_k + \mathbf{J}(\mathbf{\Omega}_{k-1})),$$

where $k = 1, \ldots, N-1$. In addition, we use the boundary conditions on $\mathbf{R}_0$ and $\mathbf{R}_N$, together with the update step given by line 1 of (31) to give the last constraint,

$$0 = \log\Big(\mathbf{R}_N^{-1} \mathbf{R}_0 \exp(h\mathbf{\Omega}_0) \ldots \exp(h\mathbf{\Omega}_{N-1}) \Big),$$

where log is the logarithm map on $\mathsf{SO}(3)$.

Note that while we use the direct variational approach to obtain the discrete extremal solutions, an alternate way to obtain the discrete extremal solutions would be to use Pontryagin's maximum principle. In particular, Bonnans and Laurent-Varin [4] show that these two approaches are equivalent in the context of symplectic partitioned Runge-Kutta schemes.

At this point it should be noted that one important advantage of the manner in which we have discretized the optimal control problem is that it is $\mathsf{SO}(3)$-equivariant. This is to say that if we rotated all the boundary conditions by a fixed rotation matrix, and solved the resulting discrete optimal control problem, the solution we would obtain would simply be the rotation of the solution of the original problem. This can be seen quite clearly from the fact that the discrete problem is expressed in terms of body coordinates, both in terms of body angular velocities and body forces. In addition, the initial and final attitudes $\mathbf{R}_0$ and $\mathbf{R}_N$ only enter in the last equation as a relative rotation.

The $\mathsf{SO}(3)$-equivariance of our numerical method is desirable, since it ensures that our results do not depend on the choice of coordinate frames. This is in contrast to methods based on coordinatizing the rotation group using quaternions and Euler angles.

Each of the equations above take values in $\mathfrak{so}(3)$. Consider the Lie algebra isomorphism between $\mathbb{R}^3$ and $\mathfrak{so}(3)$ given by the hat map,

$$\mathbf{v} = (v_1, v_2, v_3) \mapsto \hat{\mathbf{v}} = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix},$$

which maps 3-vectors to 3×3 skew-symmetric matrices. In particular, we have the following identities,

$$[\hat{\mathbf{u}}, \hat{\mathbf{v}}] = (\mathbf{u} \times \mathbf{v})\hat{}, \quad \mathrm{Ad}_{\mathbf{A}} \hat{\mathbf{v}} = (\mathbf{A}\mathbf{v})\hat{}.$$

Furthermore, we identify $\mathfrak{so}(3)^*$ with $\mathbb{R}^3$ by the usual dot product, that is to say if $\mathbf{\Pi}, \mathbf{v} \in \mathbb{R}^3$, then $\langle \mathbf{\Pi}, \hat{\mathbf{v}} \rangle = \mathbf{\Pi} \cdot \mathbf{v}$. With this identification, we have that $\mathrm{Ad}^*_{\mathbf{A}^{-1}} \mathbf{\Pi} = \mathbf{A}\mathbf{\Pi}$. Using the identities above, we write the necessary conditions using matrix-vector products and cross products. Then, each of the equations can be interpreted as 3-vector valued functions, and the system of equations can be considered as a $3(2N - 3)$-vector valued function, which is precisely the dimensionality of the unknowns. This reduces the discrete optimal problem to a nonlinear root finding problem.

The nonlinear system of equations was solved in MATLAB using the `fsolve` routine, where the Jacobian is constructed column by column, and the k-th column is computed using the following approximation [19],

$$\frac{\partial \mathbf{F}}{\partial x_k}(\mathbf{x}) = \frac{1}{\epsilon} \mathrm{Im}[\mathbf{F}(\mathbf{x} + i\epsilon \mathbf{e}_k)],$$

where $i = \sqrt{-1}$, $\mathbf{e}_k$ is a basis vector in the x_k direction, and ϵ is of the order of machine epsilon. This method is preferable to a finite-difference approximation, since it does not suffer from round-off errors, which would otherwise limit how small ϵ can be.

In our numerical simulation, we computed an optimal trajectory for a rest-to-rest maneuver, as illustrated in Figure (2). Here, the maneuver time is 12.8sec, $N = 128$, and the moment of inertia is given by

$$\mathbf{J} = \begin{bmatrix} 13.25 & -7.80 & -11.40 \\ -7.80 & 16.25 & 4.71 \\ -11.40 & 4.71 & 18.37 \end{bmatrix}.$$

The prescribed maneuver corresponds to a rotation by $\frac{\pi}{3}$ about the x-axis. Since the moment of inertia tensor is not a multiple of the identity, and the x-axis does not correspond to the axis of minimal inertia, the optimal trajectory does not just involve a pure rotation about the x-axis. It is worth noting that the results are not rotationally symmetric about the midpoint of the simulation interval, which is due to the fact that our choice of update, $\mathbf{R}_{k+1} = \mathbf{R}_k \exp(h\mathbf{\Omega}_k)$, does not exhibit time-reversal symmetry. In a forthcoming publication, we will introduce a reversible algorithm to address this issue. In particular, this will involve explicitly computing the stationarity conditions for the discrete optimal control problem constrained by the time-symmetric Lie Störmer–Verlet method.

We also present results for an optimal slew-up maneuver, illustrated in Figure (3). This uses the same moment of inertia tensor as in the previous simulation, and the desired maneuver involves a rotation of $\frac{\pi}{6}$ about the x-axis from rest to a final

angular velocity of $\Omega_{N-1} = \begin{bmatrix} 0.3 & 0.2 & 0.3 \end{bmatrix}^{\mathrm{T}}$, over a maneuver time of 12.8sec, and $N = 128$.

6. Conclusion

In this paper we studied the continuous- and discrete-time optimal control problem for the rigid body, where the cost to be minimized is the external torque applied to move the rigid body from an initial condition to some pre-specified terminal condition. In the discrete setting, we use the discrete Lagrange–d'Alembert principle to obtain the discrete equations of motion. The kinematics were discretized to guarantee that the flow in phase space remains on the Lie group $\mathsf{SO}(3)$ and its algebra $\mathfrak{so}(3)$. We described how the necessary conditions can be solved for the general three-dimensional case and gave a numerical example for a three-dimensional rigid body maneuver.

The synthesis of variational mechanics with discrete-time optimal control is particularly advantageous from the point of view of computational efficiency, since the symplectic Euler method is symplectically conjugate to the Störmer–Verlet method, and hence has effective order two. Consequently, for our discrete-time optimal control method, the cost functional converges at a rate which is characteristic of a second-order method, while being based on a first-order method that is computationally cheaper.

Currently, we are investigating the use of the Pontryagin's maximum principle with Lie group methods in continuous- and discrete-time to obtain the necessary conditions. Additionally, we wish to generalize the result to general Lie groups that have applications other than the rigid body motion on $\mathsf{SO}(3)$. In particular, we are interested in controlling the motion of a rigid body *in space*, which corresponds to motion on the non-compact Lie group $\mathsf{SE}(3)$.

References

[1] New horizons: Pluto-kuiper belt mission. Website (2006). URL http://www.nasa.gov/mission_pages/newhorizons/main/index.html

[2] Agrachev, A.A., Sachkov, Y.: Control Theory from the Geometric Viewpoint. Springer-Verlag, New York, NY (2004)

[3] Bloch, A.: Nonholonomic Mechanics and Control. Springer-Verlag, New York, NY (2003)

[4] Bonnans, J., Laurent-Varin, J.: Computation of order conditions for symplectic partitioned runge-kutta schemes with application to optimal control. Num. Math. **103**, 1–10 (2006)

[5] Hairer, E., Lubich, C., Wanner, G.: Geometric Numerical Integration. Springer, Berlin (2002)

[6] Hairer, E., Lubich, C., Wanner, G.: Geometric numerical integration illustrated by the Störmer-Verlet method. Acta Numer. **12**, 399–450 (2003)

[7] Hussein, I.I., Bloch, A.M.: Optimal control on Riemannian manifolds with potential fields. 43rd IEEE Conference on Decision and Control (2004). 1982–1987

[8] Hussein, I.I., Bloch, A.M.: Optimal trajectory tracking on the group of rigid body motions. 2005 IEEE Conference on Decision and Control (2005)

[9] Hussein, I.I., Scheeres, D.J., Hyland, D.C.: Interferometric observatories in Earth orbit. Journal of Guidance, Control and Dynamics **27**(2), 297–301 (2004)

[10] Iserles, A., Munthe-Kaas, H., Nørsett, S.P., Zanna, A.: Lie group methods. Acta Numerica **9**, 215–265 (2000)

[11] Junge, O., Marsden, J.E., Ober-Blöbaum, S.: Discrete mechanics and optimal control. IFAC Congress, Praha (2005)

[12] Kane, C., Marsden, J.E., Ortiz, M., West, M.: Variational integrators and the newmark algorithm for conservative and dissipative mechanical systems. International Journal of Numerical Methods in Engineering **49**(10), 1295–1325 (2000)

[13] Khaneja, N., Glaser, S.J., Brockett, R.W.: Sub-Riemannian geometry and optimal control of three spin systems. Physical Review A **65**, 032,301 (2002)

[14] Lee, T., Leok, M., McClamroch, N.H.: A Lie group variational integrator for the attitude dynamics of a rigid body with applications to the 3D pendulum. Proceedings of the IEEE Conference on Control Applications pp. 962–967 (2005)

[15] Lee, T., Leok, M., McClamroch, N.H.: Lie group variational integrators for the full body problem. Computer Methods in Applied Mechanics and Engineering (2005). Submitted, arXiv:math.NA/0508365

[16] Leimkuhler, B., Reich, S.: Simulating Hamiltonian Dynamics, *Cambridge Monographs on Applied and Computational Mathematics*, vol. 14. Cambridge University Press, Cambridge (2004)

[17] Leok, M.: Generalized galerkin variational integrators (2004). Preprint, arXiv:math.NA/0508360

[18] Littell, T., Skeel, R., Zhang, M.: Error analysis of symplectic multiple time stepping. SIAM J. Numer. Anal. **34**(5), 1792–1807 (1997)

[19] Lyness, J., Moler, C.: Numerical differentiation of analytic functions. SIAM J. Numer. Anal. **4**, 202–210 (1967)

[20] Marsden, J., West, M.: Discrete mechanics and variational integrators. Acta Numerica **10**, 357–514 (2001)

[21] Marsden, J.E., Ratiu, T.S.: Introduction to Mechanics and Symmetry. Springer-Verlag, New York, NY (1999)

[22] Milnor, J.: Morse Theory. Princeton University Press, Princeton, NJ (1963)

[23] Palao, J.P., Kosloff, R.: Quantum computing by an optimal control algorithm for unitary transformations. Physical Review Letters **89**, 188,301 (2002)

[24] Sanz-Serna, J.M., Calvo, M.P.: Numerical Hamiltonian Problems, *Applied Mathematics and Mathematical Computation*, vol. 7. Chapman and Hall, London (1994)

[25] Schaub, H., Junkins, J.L., Robinett, R.D.: New attitude penalty functions for spacecraft optimal control problems. AIAA Guidance, Navigation, and Control Conference (1996)

[26] Scrivener, S.L., Thompson, R.C.: Survey of time-optimal attitude maneuvers. Journal of Guidance, Control, and Dynamics **17**(2), 225–233 (1994)

[27] Suzuki, M.: Improved Trotter-like formula. Phys. Lett. A **180**(3), 232–234 (1993)

[28] Warner, F.: Foundations of Differentiable Manifolds and Lie Groups. Scott, Foresman, Glenview, IL (1973)

ALEXANDER ZIWET COLLEGIATE PROFESSOR OF MATHEMATICS AND DEPARTMENT CHAIR, MATHEMATICS, UNIVERSITY OF MICHIGAN
  E-mail address: abloch@umich.edu

ASSISTANT PROFESSOR, MECHANICAL ENGINEERING, WORCESTER POLYTECHNIC INSTITUTE
 *E-mail address*: ihussein@wpi.edu

ASSISTANT PROFESSOR, MATHEMATICS, PURDUE UNIVERSITY
  E-mail address: mleok@math.purdue.edu

ASSISTANT PROFESSOR, MECHANICAL ENGINEERING, UNIVERSITY OF HAWAII
  E-mail address: aksanyal@hawaii.edu
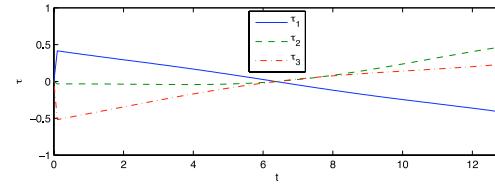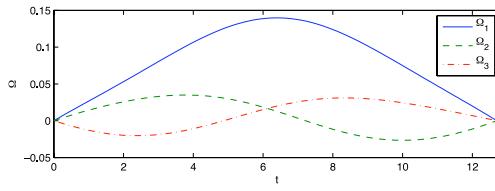
(a) Angular velocity and control torques
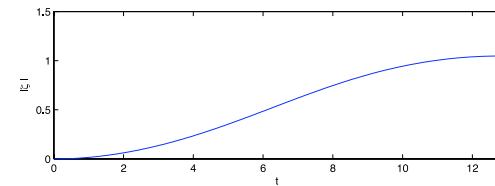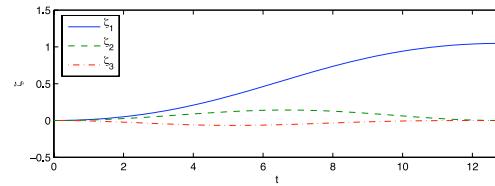

(b) Principal axis and angle
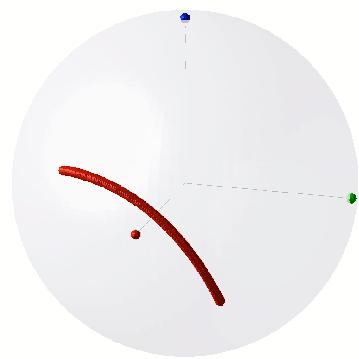
(c) Instantaneous rotation axis

FIGURE 2. Discrete optimal rest-to-rest maneuver in $\mathsf{SO}(3)$.
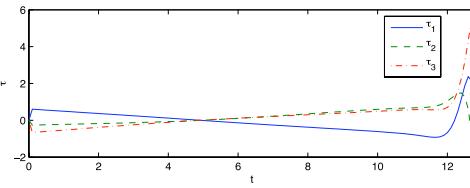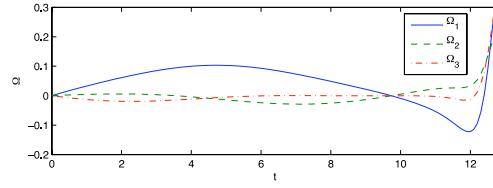
(a) Angular velocity and control torques
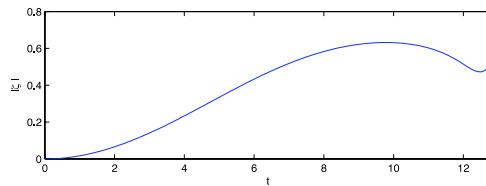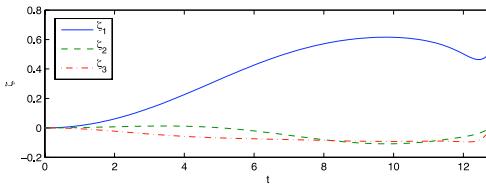
(b) Principal axis and angle
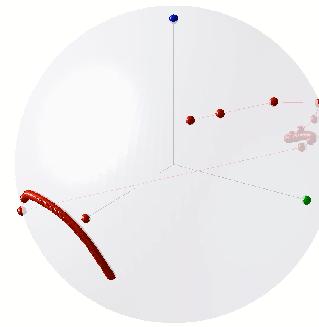
(c) Instantaneous rotation axis

FIGURE 3. Discrete optimal slew-up maneuver in $\mathsf{SO}(3)$.